

A comparative analysis of speech recognition systems for the tatar language

Khusainov A.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© Springer Nature Switzerland AG 2018. This paper presents a comparative study of several different approaches to speech recognition for the Tatar language. All the compared systems use a corpus-based approach, so recent results in speech and text corpora creation are also shown. The recognition systems differ in acoustic modelling algorithms, basic acoustic units, and language modelling techniques. The DNN-based system shows the best recognition result obtained on the test part of speech corpus.

http://dx.doi.org/10.1007/978-3-319-77113-7_40

Keywords

Acoustic modelling, Language modelling, Speech recognition, Tatar language

References

- [1] Lewis, M.P., Simons, G.F., Fennig, C.D. (eds.). Ethnologue: Languages of the World, 9th (edn.). SIL International, Dallas (2016). <http://www.ethnologue.com>
- [2] Berment, V.: "Méthodes pour informatiser des langues et des groups de langues peu dotées", Ph.D. thesis, J. Fourier University, Grenoble I (2004)
- [3] Krauwer, S.: The basic language resource kit (BLARK) as the first milestone for the language resources roadmap. In: Proceedings of International Workshop Speech and Computer SPEECOM, Moscow, Russia, pp. 8-15 (2003)
- [4] Khusainov, A.: Tekhnologiya avtomatizatsii sozdaniya i otsenki kachestva programmnikh sredstv analiza rechi c uchetom osobennostey maloresursnykh yazikov, Ph.D. thesis, Kazan, 162 p (2014)
- [5] Salimzyanov, I., Washington, J., Tyers, F.: A free/open-source Kazakh-Tatar machine translation system. In: Proceedings of the Machine Translation Summit XIV, Nice, France (2013)
- [6] Yandex Translate. <https://translate.yandex.com/translator/Russian-Tatar>
- [7] Suleymnov, D., Gatiatullin, A., Gilmullin, R.: Lexicograficheskaya baza dannykh dlya system mashinnogo perevoda blizkorodstvennykh yazikov. In: Proceedings of Third International Conference «Informatizatsiya obschestva», Astana, Kazakhstan, pp. 585-587 (2012)
- [8] Khusainov, A., Khusainova, A.: Speech human-machine interface for the Tatar language. In: Artificial Intelligence and Natural Language Conference, FRUCT Oy, Helsinki, pp. 60-65 (2016)
- [9] Khusainov, A., Suleymanov, D.: Language identification system for the tatar language. In: Železný, M., Habernal, I., Ronzhin, A. (eds.) SPECOM 2013. LNCS (LNAI), vol. 8113, pp. 203-210. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-01931-4_27
- [10] Suleymanov, Dz., Nevzorova, O.A., Khakimov, B.: National corpus of the tatar language "Tugan Tel": structure and features of grammatical annotation. In: Proceedings of International Conference Georgian Language and modern Technology, Tbilisi, pp. 107- 108 (2013)

- [11] Povey, D., et al.: The kald speech recognition toolkit. In: Proceedings of ASRU, pp. 1-4 (2011)
- [12] Rath, S.P., Povey, D., Vesely, K., Cernocky, J.H.: Improved feature processing for deep neural networks. In: Proceedings of InterSpeech (2013)
- [13] Zhang, X., Trmal, J., Povey, D., Khudanpur, S.: Improving deep neural network acoustic models using generalized maxout networks. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2014, pp. 215-219. IEEE (2014)
- [14] Stolcke, A.: SRILM - an extensible language modeling toolkit. In: Proceedings of International Conference on Spoken Language Processing, vol. 2, Denver, pp. 901-904 (2002)
- [15] Stolcke, A.: Entropy-based pruning of backoff language models. In: Proceedings of DARPA Broadcast News Transcription and Understanding Workshop, Lansdowne, pp. 270-274 (1998)
- [16] Kneser, R., Ney, H.: Improved backingoff for m-gram language modeling. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1 (1995)